October 28, 2016

Equity Market Impact Modeling: an Empirical Analysis for Chinese Market

Shiyu Han[a], Lan Wu[a,1], and Yuan Cheng[a]

[a]*School of Mathematical Sciences, Peking University, Beijing 100871, China*

Market impact has become a subject of increasing concern among academics and industry experts. We put forward a price impact model which considers the heteroscedasticity of price in the time dimension and dependency between permanent impact and temporary impact. We discuss and derive the extremum of the expectation of permanent impact and realized impact by constructing several special trading trajectories. Given our use of a large trade and quote tick records of 17,213,238,343 compiled from the Chinese stock market, the model assessment ultimately suggest that our model is better than Almgren's model. Interestingly, the result of random effect analysis indicates the parameter α, which is the exponent of the impact function, is a constant with a value of around 0.7 across all stocks. Our model and empirical result would give academia some insight of mechanism of Chinese market, and can be applied to algorithm trading.
Key Words: Market impact model; Optimal impact; Model assessment; Chinese stock market.

[1]Corresponding author: lwu@pku.edu.cn

1 Introduction

A transaction cost consists of two parts, the direct cost and the indirect cost. The study of transaction cost can provide some information about the structure of market, and also some evidence about the efficiency of the market. The indirect cost is difficult to observe, measure, and control. Yet it can be optimized through trading techniques. Thus, the indirect cost is a subject of concern in academia and industry.

The former researchers have developed the basis of impact model. Basing on the model, there are a series of problems such as the optimization problem, parameters estimation and the impact functions. Some literatures also involve the application of the impact model.

[11] first decomposed price impact into permanent impact and temporary impact. They pointed out that the permanent impact was caused by a change in investors' expectations due to information gleaned from trading. The temporary impact, on the other hand, was caused by a short non-equilibrium in trading or the controlling of quotes by dealers, and it vanished after a period of time. [12] introduced the linear impact function, while [19] pointed that the impact function should be the power function. [10] proved that the permanent impact function should be linear, i.e., the power index must equal to 1 or it contradicted the non-arbitrage assumption. Moreover, the permanent impact function should be independent of the length of trading period. The temporary impact function, however, could be nonlinear and conformed more closely to microstructure theory and the empirical standard.

[1] obtained the optimal trading strategy given the conditions of the linear impact function. They also established the efficient frontier for the mean and variance of the optimal trade trajectories with different levels of risk aversion, and identified the closed-form solutions of the corresponding optimal trading strategies. [3] extended the linear temporary impact function to the power function. [2] used the results attained by [10] to estimate the impact function based on data equity at Citigroup. They showed that permanent impact cost was irrelative to the trading duration and manner, while temporary impact cost was closely related to them. Moreover, the estimation of the power exponent for permanent impact was close to 0.8, while that of temporary impact was close to 0.6.

On the convexity of the impact function, [15] and [9] pointed out that the it was a concave function of the absolute value of the transaction volume. Unlike [2], who assumed that the permanent impact function was linear, they had assumed that the temporary impact function was strictly concave. [17] found that the impact function was convex while researching the German security market. [4] and [18] showed that the impact function was a nonlinear S-form. [6] and [5] studied the impact caused by different initiatives of buying and selling.

[5] considered how factors other than changes in price influenced trading. [1] attempted to calculate the liquidity-adjusted VaR and adapted it to the context of risk management. [14] considered dynamic portfolio management based on impact cost. [7] converted the optimal execution trajectory problem with explicit and impact costs into an optimal stochastic control problem with a restricted condition.

Our motivation is to study the feature of price impact in Chinese stock market. We suppose that there exists a consistent impact function in a specific market, here we focus on Chinese market. We improve the previous impact models developed in [1] and introduce a statistical inference to assess the model. Specifically, we propose a more explicit form of the permanent impact function and temporary impact function and give the joint distribution of the statistics of the permanent impact and the temporary impact. Consequently, we acquire a strong explanation of Chinese market impact. Further, instead of looking for the optimal executing trajectory to minimize the utility function in [1](the mean-variance utility), we derive the extremum of the expectation of permanent impact and realized impact by constructing corresponding trading trajectories. In empirical analysis, we calibrate the parameters in the impact models and made a model assessment on whole data in Chinese market. The empirical process and results should have a very strong practical meaning.

In Section 2, we present some preliminary work. In Section 3, we derive our main result. In Section 4, we describe our data. Section 5 is our empirical research. The paper ends up with a conclusion in Section 6.

2 Preliminary

In this section, we set up the process of the asset price and quantify the permanent impact and the temporary impact. For the sake of convenience, most of the variables and assumptions in our model follow from the framework in [2].

2.1 Variables

Volume time

In the real market, the volatility of the asset price varies with the time over the course of one trading day. More specifically, an identically sized order can cause greater impact in the middle of the day than at the opening or closing of the trading day. To describe this phenomenon, we use volume time instead of physical time in our model. The mapping from real time to volume time can be defined as

$$\kappa(t) = t_{open} + V(t)\frac{t_{close} - t_{close}}{V}, \quad t \in [t_{open}, t_{close}], \tag{1}$$

where V(t) is the transaction volume up to time t, V is the total transaction volume over the entire trading day, and t_{open} and t_{close} represent the opening and closing times of a trading day, respectively. In the following context, if time is not specified as physical time, it should be assumed the volume time.

Price variables

As introduced in [2], the impact caused by executing an order contained a temporary compoenent, which was caused by a transient liquidity drought or other reasons and vanished after a period of time with the liquidity replenished. Here we define the beginning of the execution of an order as time 0, the end of the execution as time T, and the time at which the temporary impact vanishes as T_{post}, which is greater than T. Correspondingly, we define the prices at different times as

S_0	the price before the impact occurs
S_{post}	the price when the temporary impact vanishes
$\overline{S}$	the volume weighted average of the realized price

S_0 represents the price before the impact occurs, S_{post} represents the price when the temporary impact vanishes, and $\overline{S}$ indicates the volume weighted volume of the realized price in the execution period $[0, T]$.

Impact statistics

We decompose the impact into two components, the permanent impact and the temporary impact. The permanent impact reflects the information released during the trading process. The temporary impact reflects the loss to the counterparties caused by positive trading. The permanent impact is lasting, while the temporary impact vanishes as time passes. Two statistics to describe these two impacts are defined as follows:

$$Permannent\ impact: \quad I = \frac{S_{post} - S_0}{S_0}, \tag{2}$$

$$Realized\ impact: \quad J = \frac{\overline{S} - S_0}{S_0}. \tag{3}$$

According to these two impact statistics, the temporary impact can be expressed as $I - J/2$.

2.2 Price Process

We assume asset price S_t follows the arithmetic Brownian motion,

$$dS_t = S_0 g(v_t)dt + S_0 \sigma dB_t, \quad t \geq 0, \tag{4}$$

where the trading velocity v_t is deterministic, $g(v_t)$ in the drift term is the permanent impact function, which depends on trading velocity v_t and $g(v_t) = 0$ when $t > T$, and B_t is the standard Brownian motion. It is important to point out that asset price S_t in (4) is the real price but not the transaction price, which we denote by $\tilde{S}_t$ in the following assumption:

$$\tilde{S}_t = S_t + S_0 h(v_t), \tag{5}$$

where $h(v)$ is the temporary impact function and depends on the trading velocity as well.

Under the process assumption above, we can calculate the concrete form of (2) and (3) as follows (a detailed calculation is provided in the appendix):

$$I = \int_0^T g(v_t)dt + \sigma B_{T_{post}}, \tag{6}$$

$$J = \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt + \frac{\sigma}{T}\int_0^T B_t dt. \tag{7}$$

The impact functions $g(\cdot)$ and $h(\cdot)$ in (6) and (7) have been studied by various researchers, as discussed in the introduction. In our work, we adopt the power functions as previous literatures:

$$g(v) = \gamma sgn(v)|v|^{\alpha}, \tag{8}$$

$$h(v) = \eta sgn(v)|v|^{\beta}, \tag{9}$$

3 Main Result

3.1 The joint distribution of I and J

At first, we derive the joint distribution of I and J as the following proposition.

Propostion 1. *I and J defined in (6) and (7) follow Bivariate Normal Distribution as $(I, J)^T \sim N(\mu, \Sigma)$, where*

$$\mu = \begin{pmatrix} \int_0^T g(v_t)dt \\ \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma^2 T_{post} & \frac{1}{2}\sigma^2 T \\ \frac{1}{2}\sigma^2 T & \frac{1}{3}\sigma^2 T \end{pmatrix}.$$

Proof. According to the property of Gaussian process, I and J follow Bivariate Normal Distribution. Basing on (6) and (7), we can easily obtain the covariance matrix

by the property of Ito integral.

$$
\begin{aligned}
Var(I) = E[I - E(I)]^2 &= E(\sigma^2 B_{T_{post}}^2) = \sigma^2 T_{post}, \\
Var(J) = E[J - E(J)]^2 &= E[\frac{\sigma}{T}\int_0^T (T-t)dB_t]^2 \\
&= \frac{\sigma^2}{T^2}\int_0^T (T-t)^2 dt \\
&= \frac{1}{3}\sigma^2 T.
\end{aligned}
$$

The covariance of I and J can be calculated as follows:

$$
\begin{aligned}
Cov(I,J) &= E[(I - E(I))(J - E(J))] \\
&= E[\sigma \int_0^{T_{post}} dB_t \cdot \frac{\sigma}{T}\int_0^T (T-t)dB_t] \\
&= \frac{\sigma^2}{T} E[\int_0^T dB_t \int_0^T (T-t)dB_t] \\
&= \frac{\sigma^2}{T}\int_0^T (T-t)dt \\
&= \frac{1}{2}\sigma^2 T.
\end{aligned}
$$

Therefore, I and J follow the bivariate normal distribution as $(I,J)^T \sim N(\mu, \Sigma)$, where

$$
\mu = \begin{pmatrix} \int_0^T g(v_t)dt \\ \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma^2 T_{post} & \frac{1}{2}\sigma^2 T \\ \frac{1}{2}\sigma^2 T & \frac{1}{3}\sigma^2 T \end{pmatrix}.
$$

□

3.2 Optimal Impact

Impact always exists in the trading process, though the magnitude varies. Instead of attempting to describe or even control for impact cost in the trading process, market participants may pay attention to the description of the impact directly. More

specifically, in terms of expectations, market participants may ask: What is the optimum of the impact for a transaction of a fixed number of shares over a specified time interval? In what follows, we provide a specific solution of this problem given the explicit form of the impact function.

As mentioned above, we first suppose there are X shares to be transacted over a given time interval $[0, T]$ (note that the time here is volume time) in a specific direction, i.e., buy or sell. Note that we want to find the optimum of the impact by transaction. In the following discussion, we denote the trajectory of trading by x_t, with an initial position of $x_0 = 0$ and a final position of $x_T = X$, which keeps a buy direction. Furthermore, the trading velocity v_t does not change the sign. Similarly, we can analysis the opposite direction of sell.

Following the previous part of the paper, we take I as the permanent impact, which exists even after the final transaction, while J is the realized impact, which is the average impact during the trading process. Therefore, we can formulate the best trading trajectory to maximize or minimize the expected impact.

We propose the following optimizations:

$$\begin{cases} \min \ or \ \max \ \ E[I] \\ s.t. \quad x_0 = 0, \ x_T = X \ and \ v_t \geq 0, \end{cases} \tag{10}$$

and

$$\begin{cases} \min \ or \ \max \ \ E[J] \\ s.t. \quad x_0 = 0, \ x_T = X \ and \ v_t \geq 0, \end{cases} \tag{11}$$

where $v_t = \dot{x}_t$.

The optimization problems (10) and (11) can be generalized to the following variational problems:

$$\min \ or \ \max \int_0^T F(t, x_t, v_t) dt. \tag{12}$$

In the permanent impact case, we have

$$F(t, x_t, v_t) = g(v_t), \tag{13}$$

in the realized impact case, we have

$$F(t, x_t, v_t) = \frac{T-t}{T} g(v_t) + \frac{1}{T} h(v_t). \tag{14}$$

To solve problem (12), we consider the Euler-Lagrange equation:

$$F_x(t, x, v) = \frac{d}{dt} F_v(t, x, v). \tag{15}$$

By substituting (13) and (14) into the Euler-Lagrange equation respectively, we obtain two differential equations. In (13) case, the equation is:

$$\gamma v \alpha(\alpha - 1) v^{\alpha-2} = 0, \tag{16}$$

while in (14) case, the equation is:

$$-\frac{1}{T}\gamma\alpha v^{\alpha-1} + \frac{T-t}{T}\gamma\alpha(\alpha-1)v^{\alpha-2}\dot{v} + \frac{1}{T}\eta\beta(\beta-1)v^{\beta-2}\dot{v} = 0. \tag{17}$$

The optimization problems (10) and (11) can be solved through equation (16) and equation (17) with the boundary conditions.

Equations (16) and (17) are mainly parameterized by α and β, and they involve nonlinear problems for which no closed-form solutions can be obtained. Nevertheless, we can still obtain analytical results in special cases, which are shown in what follows.

Propostion 2. *The solutions of optimization problem (10) and (11) for some specific case are as bellow,*

(1) For problem (10)
when $\alpha < 1$, $\min E(I) = 0$, $\max E(I) = \gamma T^{1-\alpha} X^\alpha$,
when $\alpha = 1$, $\min E(I) = \max E(I) = \gamma X$,
when $\alpha > 1$, $\min E(I) = \gamma T^{1-\alpha} X^\alpha$, $\max E(I) = +\infty$.

(2) For problem (11)

when $\alpha = 1, \beta < 1, \beta \neq \frac{1}{2}$,

$\min E(J) = 0$,

$\max E(J) = \frac{\eta^2\beta^2(\beta-1)}{\gamma T(2\beta-1)}[(\frac{\gamma}{\eta\beta}T + C_1)^{\frac{2\beta-1}{\beta-1}} - C_1^{\frac{2\beta-1}{\beta-1}}] + \gamma C_2$,

when $\alpha = 1, \beta = \frac{1}{2}$,

$\min E(J) = 0$,

$\max E(J) = \frac{\eta^2}{4\gamma T}\log(1 + \frac{2\gamma T}{\eta C_1})$,

when $\alpha = 1, \beta = 1$,

$\min E(J) = \frac{\eta X}{T}$,

$\max E(J) = \frac{\eta X}{T} + \gamma X$,

when $\alpha = 1, \beta > 1$,

$\min E(J) = \frac{\eta^2\beta^2(\beta-1)}{\gamma T(2\beta-1)}[(\frac{\gamma}{\eta\beta}T + C_1)^{\frac{2\beta-1}{\beta-1}} - C_1^{\frac{2\beta-1}{\beta-1}}] + \gamma C_2$,

$\max E(J) = +\infty$,

when $\beta = 1, \alpha < 1$,

$\min E(J) = \frac{\eta X}{T}$,

$\max E(J) = \gamma(\frac{\alpha-2}{\alpha-1})^{\alpha-1}X^\alpha T^{1-\alpha} + \frac{\eta X}{T}$,

when $\beta = 1, 1 < \alpha \leq 2$,

$\min E(J) = \frac{\eta X}{T}$,

$\max E(J) = +\infty$,

when $\beta = 1, \alpha > 2$,

$\min E(J) = \gamma(\frac{\alpha-2}{\alpha-1})^{\alpha-1}X^\alpha T^{1-\alpha} + \frac{\eta X}{T}$,

$\max E(J) = +\infty$.

where C_1, C_2 *satisfy the following equations,*

$$\frac{\eta(\beta-1)}{\gamma}(\frac{\gamma}{\eta\beta}T + C_1)^{\frac{\beta}{\beta-1}} + C_2 = X,$$
$$\frac{\eta(\beta-1)}{\gamma}C_1^{\frac{\beta}{\beta-1}} + C_2 = 0.$$

Proof. In order to determine whether a solution generated by the Euler-Lagrange equation is a maximal or a minimal solution, we construct three series of trading trajectories.

Given a $m > 0$, the first trajectory is

$$\begin{cases} x_t = \frac{X}{T^m} t^m, \\ v_t = \frac{mX}{T^m} t^{m-1}, \end{cases} \qquad t \in [0, T], \tag{18}$$

which is denoted as TA.

Given a $m > 0$, the second trajectory is

$$\begin{cases} x_t = X - \frac{X}{T^m}(T-t)^m, \\ v_t = \frac{mX}{T^m}(T-t)^{m-1}, \end{cases} \qquad t \in [0, T], \tag{19}$$

which is denoted as TB.

And given a $m > 0$, the third trajectory is

$$\begin{cases} x_t = \frac{X}{\log \frac{T+m}{m}} \log \frac{T+m}{T-t+m}, \\ v_t = \frac{X}{\log \frac{T+m}{m}} \frac{1}{T-t+m}, \end{cases} \qquad t \in [0, T], \tag{20}$$

which is denoted as TC.

Note that these three trading trajectories all satisfy the boundary condition so that $x_0 = 0$, $x_T = X$ and $v_t \geq 0$.

(1) For problem (10)

(a) $\alpha \neq 1$

In this case, equation (16) has solution $v = const > 0$ and therefore $v = X/T$. The expected permanent impact is

$$E[I] = \gamma \int_0^T (\frac{X}{T})^\alpha dt = \gamma T^{1-\alpha} X^\alpha. \tag{21}$$

Note that with the trading trajectory TA of (18), we have

$$E[I] = \frac{\gamma X^\alpha T^{1-\alpha} m^\alpha}{\alpha m - \alpha} \xrightarrow{m \to \infty} \begin{cases} \infty & \alpha > 1, \\ 0 & \alpha < 1. \end{cases}$$

Hence when $\alpha > 1$, the solution of the differential equation (16) gets a minimal expected impact, the lower bound of the permanent impact is $\gamma T^{1-\alpha} X^{\alpha}$, and the upper bound is ∞. Similarly, when $\alpha < 1$, the lower bound is 0 and the upper bound is $\gamma T^{1-\alpha} X^{\alpha}$.

(b) $\alpha = 1$

In this case, equation (16) degenerates, and the permanent impact function becomes $g(v) = \gamma v$. Accordingly, the expected permanent impact is

$$E[I] = \int_0^T \gamma v dt = \gamma X = const, \tag{22}$$

which means that any trading trajectory has the same expected permanent impact, i.e., the upper bound and lower bound are both γX.

We remark that, with a total X shares to be transacted, when $\alpha = 1$, the permanent impact is a linear pattern and consequently the expected permanent impact is a constant regardless of trading velocity. When $\alpha < 1$, a trading trajectory with uniform velocity leads to the maximal expected permanent impact. This means that one can obtain a lower expected permanent impact as long as the shares are transacted over a shorter time interval. When $\alpha > 1$, similarly, a trading trajectory with uniform velocity leads to the minimal expected permanent impact, and a shorter time interval for the transaction leads to a greater expected permanent impact.

(2) For problem (11)

(a) $\alpha = 1$ **and** $\beta \neq 1$

In this case, equation (17) reduces to

$$\gamma = \eta\beta(\beta - 1)v^{\beta-2}\dot{v}. \tag{23}$$

The closed-form solution of (23) is

$$v_t = (\frac{\gamma t}{\eta\beta} + C_1)^{\frac{1}{\beta-1}}, \tag{24}$$

$$x_t = \frac{(\beta-1)\eta}{\gamma}(\frac{\gamma t}{\eta\beta} + C_1)^{\frac{\beta}{\beta-1}} + C_2, \tag{25}$$

where C_1 and C_2 satisfy the following equations:

$$\frac{\eta(\beta-1)}{\gamma}(\frac{\gamma}{\eta\beta}T + C_1)^{\frac{\beta}{\beta-1}} + C_2 = X, \tag{26}$$

$$\frac{\eta(\beta-1)}{\gamma}C_1^{\frac{\beta}{\beta-1}} + C_2 = 0. \tag{27}$$

When $\beta \neq \frac{1}{2}$,

$$E[J] = \frac{\eta^2\beta^2(\beta-1)}{\gamma T(2\beta-1)}[(\frac{\gamma}{\eta\beta}T + C_1)^{\frac{2\beta-1}{\beta-1}} - C_1^{\frac{2\beta-1}{\beta-1}}] + \gamma C_2. \tag{28}$$

When $\beta = \frac{1}{2}$,

$$E[J] = \frac{\eta^2}{4\gamma T}\log(1 + \frac{2\gamma T}{\eta C_1}). \tag{29}$$

To verify the solution of differential equation (23) is indeed corresponding a maximal or minimal realized impact, we choose the trading trajectory TA of (18), the expected realized impact of which is

$$E[J] = \frac{\gamma}{m+1}X + \frac{\gamma X^\beta T^{-\beta} m^\beta}{\beta m - \beta + 1} \xrightarrow{m\to\infty} \begin{cases} \infty, & \beta > 1, \\ 0, & \beta < 1. \end{cases} \tag{30}$$

Hence, when $\beta < 1$, we conclude that the solution of differential equation (23) corresponds to the maximal expected realized impact. When $\beta > 1$, we obtain the minimal expected realized impact. More specifically, when $\beta \neq 1/2$ and $\beta < 1$, the lower bound of the expected realized impact is 0 and the upper bound is (28). When $\beta = 1/2$, the lower bound is 0 and the upper bound is (29). On the other hand, when $\beta > 1$, the lower bound is (28) and the upper bound is ∞.

(b) $\alpha \neq 1$ **and** $\beta = 1$

In this case, equation (17) reduces to

$$\frac{1}{T}\gamma\alpha v^{\alpha-1} = \frac{T-t}{T}\gamma\alpha(\alpha-1)v^{\alpha-2}\dot{v}. \tag{31}$$

When $\alpha > 2$ or $\alpha < 1$, the closed-form solution that corresponds to this case is

$$v_t = \frac{X}{T}(T-t)^{\frac{1}{1-\alpha}}, \tag{32}$$

$$x_t = X[1-(\frac{T-t}{T})^{\frac{\alpha-2}{\alpha-1}}]. \tag{33}$$

and when $1 < \alpha \leq 2$, there is no solution to the differential equation.

Note that when $\beta = 1$, the expected realized impact is

$$\begin{aligned} E[J] &= \int_0^T \frac{T-t}{T}\gamma v^\alpha dt + \frac{1}{T}\int_0^T \eta v dt \\ &= \frac{\gamma}{T}\int_0^T (T-t)v^\alpha dt + \frac{\eta}{T}X \\ &\geq \frac{\eta X}{T}, \end{aligned}$$

and the solution of (31) has expected realized impact

$$E[J] = \gamma(\frac{\alpha-2}{\alpha-1})^{\alpha-1}X^\alpha T^{1-\alpha} + \frac{\eta X}{T}. \tag{34}$$

In order to verify whether (34) has the maximal or minimal expected realized impact, we also resort to trading trajectories TA of (18) and TB of (19). When $\alpha < 2$, the expected realized impact of trading trajectory TA of (18) is

$$E[J] = \frac{\gamma m^\alpha X^\alpha T^{1-\alpha}}{(\alpha m - \alpha + 1)(\alpha m - \alpha + 2)} + \frac{\eta X}{T} \xrightarrow{m\to\infty} \frac{\eta X}{T}. \tag{35}$$

When $\alpha > 1$, the expected realized impact of trading trajectory TB of (19)

is

$$E[J] = \frac{\gamma m^{\alpha} X^{\alpha} T^{1-\alpha}}{\alpha m - \alpha + 2} + \frac{\eta X}{T} \xrightarrow{m\to\infty} \infty. \tag{36}$$

This indicates that when $\alpha < 1$, with a large enough m, trading trajectory TA of (18) has a lower expected realized impact than the solution of differential equation (31). Hence the solution results in the maximal expected realized impact, i.e., the lower bound is $\eta X/T$ and the upper bound is (34).

Similarly, when $\alpha > 2$, the solution of the differential equation results in the minimal expected realized impact, i.e., the lower bound is (34) while the upper bound is ∞. When $1 < \alpha < 2$, there are two sequences of trading trajectories that have a realized impact converging to $\eta X/T$ and ∞, respectively. This implies that the realized impact interval is $(\eta X/T, \infty)$.

When $\alpha = 2$, the expected realized impact of trading trajectory TC of (20) is

$$E[J] = \frac{\gamma X^2}{T \log^2 \frac{T+m}{m}} \int_0^T \frac{T-t}{(T-t+m)^2} dt + \frac{\eta X}{T} \tag{37}$$

$$= \frac{\gamma X^2}{T} \frac{\log \frac{T+m}{m} - \frac{T}{T+m}}{\log^2 \frac{T+m}{m}} + \frac{\eta X}{T} \xrightarrow{m\to 0+} \frac{\eta X}{T} \tag{38}$$

Combined with trading trajectory TB of (19), we can conclude that the realized impact interval is $(\eta X/T, \infty)$.

(c) $\alpha = 1$ **and** $\beta = 1$

In this case, equation (17) reduces to

$$\frac{\gamma}{T} = 0. \tag{39}$$

Hence this condition results in no solution to differential equation (17).

Indeed, note that the expected realized impact in this case is

$$
\begin{aligned}
E[J] &= \int_0^T \frac{T-t}{T}\gamma v dt + \frac{1}{T}\int_0^T \eta v dt \\
&= \frac{\gamma}{T}\int_0^T x_t dt + \frac{\eta X}{T} \in (\frac{\eta X}{T}, \frac{\eta X}{T} + \gamma X).
\end{aligned}
$$

Similar to the former cases, trading trajectory TA of (18) has an expected realized impact of

$$
E[J] = \frac{\gamma}{m+1}X + \frac{\eta}{T}X \xrightarrow{m\to\infty} \frac{\eta X}{T}, \tag{40}
$$

and trading trajectory TB of (19) corresponds to

$$
E[J] = \gamma X - \frac{\gamma X}{m+1} + \frac{\eta X}{T} \xrightarrow{m\to\infty} \frac{\eta X}{T} + \gamma X. \tag{41}
$$

This means that no trading trajectory can result in the maximal or minimal expected realized impact, i.e., the impact interval is $(\eta X/T, \eta X/T + \gamma X)$, which is consistent with the result that the the Euler-Lagrange equation has no solution.

(d) Other cases

In addition to the above cases, the ODE (17) is general nonlinear and its solvability remains more specific discussion.

The summary remark of above are that, with a fixed X, when $\beta = 1$, the temporary impact is a linear pattern. Similar to the permanent impact case, the temporary impact is invariant regardless of trading velocity. Then the realized impact is only affected by the permanent impact. When $\alpha < 1$, a trading trajectory with a decreasing velocity gets a maximal permanent impact, which is positively related to T. When $\alpha = 1$, the permanent impact is independent of T. When $1 < \alpha \leq 2$, the permanent impact can be any positive value regardless of T. When $\alpha > 2$, a trading trajectory with increasing velocity gets a minimal permanent impact, which is negatively related to T.

□

The Propostion 2 implies that we can control the permanent impact and the realized impact when α is not equal to 1. This means trader can manipulate the underlying price to some extent. By quickly buying a large volume of shares and then slowly selling part of the previously bought shares, she can open the position with nearly no price impact. By repeating this procedure, she can set up a large enough position with a small price impact. She buys slowly to push the price high enough, then she closes her position with a small impact in a way opposite to her position direction.

Finally trader can profit just because the execute manner, which is a evidence of the inefficiency of the market. The farther that α deviates from 1, the more inefficient the market is and the greater the opportunity for this kind of profiting is.

4 Data description

Our empirical work is based on Chinese stock market data from January 2006 to January 2016. Initially we chose 2,542 stocks, giving us a corresponding total number of trade and quote (TAQ) tick records of 17,213,238,343. Taking into account that some stocks may have had invalid or incomplete data, we filtered the data sample. Ultimately, we selected 1,993 stocks that had valid data for at least five years, making the total number of tick records for the chosen stocks 16,189,238,181.

Note that we only had the tick-by-tick data; we did not have the true orders submitted by all traders. Therefore, we adopted the algorithm introduced by [13] to identify the orders. The Lee-Ready algorithm suggested that we could identify the order as buyer-initiated if the transaction price was closer to the prevailing best ask price, and we could identify it as seller-initiated if the transaction price was closer to the prevailing best bid price. If the order was transacted at the prevailing mid-quote price, we identified it as buyer-initiated if the transacted price was higher than the prevailing transacted price, and seller-initiated if the transacted price was lower than the prevailing transacted price.

5 Empirical analysis

We apply the impact model to Chinese stock market to explore the mechanism of Chinese market. Specifically, we establish the statistical model of the price impact. Moreover, we assess our model by using some criteria including BIC and CRPS.

5.1 Statistical model

When estimating the parameters of the impact functions, [2] did not consider the correlation between the temporary impact and the permanent impact. Instead, we calibrate the parameters based on the joint distribution of the impact statistics. Furthermore, [2] only considered the heteroskedasticity across the underlying, while we consider the heteroskedasticity of volatility in the time dimension besides that.

To estimate the parameters in our model, it is necessary to address the discrete version of our model based on our observed variables. We take 15 minutes as the time unit and analyse the impact caused by the executing of orders within each 15-minute segment. In the following of this part, the time tick t denotes the number of the time unit of trading period. In addition, we choose t_{post} as 15 minutes after the executing time interval.

In the discrete version of our model, we denote $(I_t, J_t)^T$ by K_t. And v_t, V_t, σ_t are the trading velocity, trading volume, and volatility, respectively, within 15 minutes. In addition, v_t is held to be positive if the trading direction is buying, and negative if it is selling. Note that the subindex t of σ_t above reflects the heteroskedasticity of volatility in the time dimension. We suppose that σ_t explains the heteroskedasticity better than variables like turnover. The general way to calculate σ_t is using the last price. But for the case of high frequency data, the return has negative correlation due to the bid-ask spread. As a result, the volatility is over-estimated. Instead we introduce a method for computation of σ_t according to the so-called smart-price, which is a better estimate for the real price of the underlying in 15 minutes. The smart-price is defined as,

$$smartprice_t = (P_a * Q_b + P_b * Q_a)/(Q_a + Q_b),$$

where P_a, P_b, Q_a, Q_b are the best ask price, the best bid price, and the corresponding volumes, respectively.

The distribution of K_t follows Propostion 1. Recognizing that the trading volume across stocks varies, we can eliminate this effect by normalizing trading vocality v_t by corresponding trading volume V_t. Hence our model can be expressed as

$$g(v_t) = \gamma\sigma_t sgn(v_t)(\frac{v_t}{V_t})^{\alpha}, \tag{42}$$

$$h(v_t) = \eta\sigma_t sgn(v_t)(\frac{v_t}{V_t})^{\beta}, \tag{43}$$

where $sgn(\cdot)$ is the sign function.

5.2 Model estimation

With the joint distribution of impact statistics K_t being Bivariate Normal as shown in Propostion 1, it is natural to use the maximum likelihood method to estimate the parameters of our model. The log-likelihood function of one stock over the all observations is

$$\begin{aligned}
&\mathcal{L}(\alpha,\beta,\gamma,\eta,K_t,v_t,V_t,\sigma_t,t\in\mathcal{T}) \\
&\quad = -\frac{1}{2}\sum_{t\in\mathcal{T}}\Big[(K_t-\mu_t)^T\Sigma_t^{-1}(K_t-\mu_t)+log(|\Sigma_t|)+2log(2\pi)\Big],
\end{aligned}$$

where $\mathcal{T}$ is the data set, and

$$\mu_t = \begin{pmatrix} T\gamma\sigma_t sgn(v_t)(\frac{v_t}{V_t})^{\alpha} \\ \frac{1}{2}T\gamma\sigma_t sgn(v_t)(\frac{v_t}{V_t})^{\alpha} + \eta\sigma_t sgn(v_t)(\frac{v_t}{V_t})^{\beta} \end{pmatrix},$$

$$\Sigma_t = \begin{pmatrix} \sigma_t^2 T_{post} & \frac{1}{2}\sigma_t^2 T \\ \frac{1}{2}\sigma_t^2 T & \frac{1}{3}\sigma_t^2 T \end{pmatrix}.$$

First we estimate the parameters for each stock. The mean, maximal, minimal of time ticks for each stock are 28,564, 38,249 and 18,039 respectively. And the results are summarized in Figure **??**. At first glance, the four parameters are distributed as bell-shaped curves. Figure **??** shows that α is mainly distributed in the interval

of $[0.6, 0.75]$ and β in $[0.6, 0.8]$. In addition, we can see that the distributions of the parameters are concentrated, which motivates us to explore whether the parameters are constants shared by all stocks.

We proceed to estimate the consistent parameters jointly data over all stocks as well as in different boards, including the Main Board of the Shanghai Stock Exchange, the Main Board of the Shenzhen Stock Exchange, the Small and Medium Enterprise Board, and the Growth Enterprise Market. The total amount of time ticks for these five sets are 56,927,427, 25,442,267, 12,311,959, 13,730,481, and 5,442,720 respectively. Table 1 lists the results of the estimation, with the first column providing the abbreviations for the boards and the numbers in parentheses being the standard errors. We can see that α and β are significantly greater than 0.5 yet less than 1. As for γ and η, they are significantly different. These results attribute to the different dimensions of the impact functions $g(v)$ and $h(v)$.

The parameter α we obtained in the Chinese market is systematically less than that in the American market as obtained by [2], which indicates that the Chinese market is less efficient than the American market. This difference might be due to the immaturity of Chinese investors and the market.

Table 1: The result of joint estimating for parameters.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\eta}$
All	0.6866 (0.0012)	0.7090 (0.0024)	4.5713 (0.0088)	0.0520 (0.0002)
SH	0.6799 (0.0023)	0.6986 (0.0045)	4.5796 (0.0164)	0.0547 (0.0004)
SZ	0.6827 (0.0021)	0.6769 (0.0039)	4.5321 (0.0127)	0.0537 (0.0003)
SME	0.6835 (0.0019)	0.7048 (0.0042)	4.4363 (0.0123)	0.0506 (0.0003)
GME	0.6849 (0.0028)	0.7683 (0.0075)	4.4835 (0.0149)	0.0490 (0.0003)

5.3 Model assessment

To justify our estimation results, we compare our model with the one presented in [2]. We use the continuously ranked probability score (CRPS, [16]) and the BIC criterion.

CRPS is a proper scoring rule which was used by [8] to assess the estimated probability distribution. When the estimated probability distribution coincides with the true probability distribution, the expectation for CRPS reaches its maximum,

which is defined as

$$CRPS(F, y) = -\int_{-\infty}^{\infty} (F(x) - 1_{\{y \leq x\}})^2 dx,$$

where F is the estimated cumulative distribution and y is an observation. It can also be expressed as

$$CRPS(F, y) = \frac{1}{2} E_F|Y - Y'| - E_F|Y - y|,$$

where Y and Y' are independent variables with distribution F.

We compare the CRPS of our model with that of [2]. According to its property, we expect to obtain a relatively greater CRPS. Thus, the comparision is assessed by the difference between the CRPS of our model and that of [2] across all stocks. Figure **??** (a-c) shows that the difference is mostly positive; therefore, our model outperforms the model in [2].

In addition, BIC is an often-used criterion for model selection. Note that BIC is consistent, and choosing the model with the minimum BIC is equivalent to choosing the model with the maximum post-probability. Thus we choose the model with the lower BIC. Similarly, we compare the BIC of our model and those of [2] across all stocks. Figure **??** shows that in most cases, our model has lower BIC that the model produced by [2].

Furthermore, we compare the CRPS and BIC for all stocks and the aforementioned boards. Figure 2 shows that the values of CRPS and BIC in our model are better than those of [2]. Consequently, our model is more reliable.

Table 2: The result of joint estimating for parameters.

	CRPS for I		CRPS for J		CRPS for $J - \frac{I}{2}$		BIC	
	Our	Alm2005	Our	Alm2005	Our	Alm2005	Our	Alm2005
All	-0.0277	-0.0298	-0.0061	-0.0097	-0.0127	-0.0127	1867891.18	2040472.65
SH	-0.0280	-0.0287	-0.0056	-0.0095	-0.0123	-0.0131	748337.13	1144537.64
SZ	-0.0281	-0.0301	-0.0063	-0.0101	-0.0129	-0.0130	893682.96	955813.35
SME	-0.0283	-0.0307	-0.0063	-0.0100	-0.0129	-0.0131	974733.32	1011478.44
GME	-0.0281	-0.0292	-0.0066	-0.0103	-0.0128	-0.0134	514790.13	592355.86

5.4 Random effect of α

The results of the consistent estimate of α indicate that all of the stocks in the market may share the same α. To verify this conjecture, we convert our model to a random effect model. More specifically, we assume that α is distributed as $N(\mu_\alpha, \sigma_\alpha^2)$. We introduce a hypothesis testing $H_0 : \sigma_\alpha = 0\, v.s.\, H_1 : \sigma_\alpha > 0$. If we reject H_0, we conclude α has random effect, which means that all of the stocks do not share the same α. On the contrary, we believe α is a constant.

As in previous work, we analysed the random effect based on all of the stocks and the boards. The first column in Table 3 is $\hat{\sigma}_\alpha$ for our model, while the third column is for Almgren's model. The second and fourth columns are the p-values corresponding to the two models. In our model, we cannot reject the hypothesis H_0, while the result in Almgren's model significantly rejects H_0. Hence all of the stocks in the Chinese stock market share the same α in our model. As it is more consistent, our model performs better than Almgren's model in explaining the Chinese stock market.

Table 3: Result of random effect analysis.

	$\hat{\sigma}_\alpha$	pvalue	$\hat{\sigma}_\alpha$	pvalue
ALL	0.0005	0.4776	0.0184	0.0007
SH	0.0010	0.3123	0.0325	0.0012
SZ	0.0009	0.3742	0.0278	0.0009
SME	0.0009	0.1787	0.0225	0.0010
GME	0.0012	0.3278	0.0055	0.0012

6 Conclusions

Based on the assumptions of the price process and the impact functions, we formulated the models of permanent and temporary impact. We obtained the explicit optima of the expected realized impact and the permanent impact by means of constructing three series of trading trajectories for specific cases. Our findings suggest that when α is less than 1, the market is inefficient.

We improve upon the model in [2] in two aspects. First, we take into account the heteroscedasticity in the time dimension by introducing the time-varying volatility

model for a more in line with the reality of the market. In empirical analysis, we calibrate the market consistent parameters on the joint model with correlation between the permanent impact and the temporary impact. Moreover, CRPS and BIC indicate that our model is systematically better than Almgren's model. The test of random effect suggests α is a constant with statistically significant in the market. This result justifies that our model can reveal some mechanism of Chinese stock market.

We applied our model to the Chinese stock market with a large trade and quote dataset. The result of the empirical analysis implies that the Chinese market is less efficient than the American market. The findings provide some insights of the market and have significant guidance for regulators. For instance, regulators might enhance the trade charge for short swing trading or limit short swing trading. In practice, our model can also be applied to algorithm trading.

Appendix A: Calculation of the Impact Variables

Permanent Impact

$$
\begin{aligned}
I =& \frac{S_{post} - S_0}{S_0} \\
=& \frac{S_0 + \int_0^T S_0 g(v_t)dt + \int_0^{T_{post}} S_0 \sigma dB_t - S_0}{S_0} \\
=& \int_0^T g(v_t)dt + \sigma B_{T_{post}}.
\end{aligned}
$$

Realized Impact

$$
\begin{aligned}
J &= \frac{\overline{S} - S_0}{S_0} = \frac{\frac{1}{T}\int_0^T \tilde{S}_t dt - S_0}{S_0} \\
&= \frac{\frac{1}{T}\int_0^T [S_t + S_0 h(v_t)]dt - S_0}{S_0} \\
&= \frac{1}{T}\int_0^T [(\int_0^t g(v_s)ds + \sigma B_t) + h(v_t)]dt \\
&= \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt + \frac{\sigma}{T}\int_0^T B_t dt \\
&= \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt + \sigma B_T - \frac{\sigma}{T}\int_0^T tB_t \\
&= \int_0^T \frac{T-t}{T} g(v_t)dt + \frac{1}{T}\int_0^T h(v_t)dt + \frac{\sigma}{T}\int_0^T (T-t)dB_t.
\end{aligned}
$$

References

[1] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[2] Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li. Direct estimation of equity market impact. *Risk*, 18(7):58–62, 2005.

[3] Robert F Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied mathematical finance*, 10(1):1–18, 2003.

[4] Utpal Bhattacharya and Matthew Spiegel. Insiders, outsiders, and market breakdowns. *Review of Financial Studies*, 4(2):255–282, 1991.

[5] Jacob A Bikker, Laura Spierdijk, and Pieter Jelle Van Der Sluis. Market impact costs of institutional equity trades. *Journal of International Money and Finance*, 26(6):974–1000, 2007.

[6] Louis KC Chan and Josef Lakonishok. Institutional trades and intraday stock price behavior. *Journal of Financial Economics*, 33(2):173–199, 1993.

[7] Thangaraj Draviam, Thomas F Coleman, and Yuying Li. Dynamic liquidation under market impact. *Quantitative Finance*, 11(1):69–80, 2011.

[8] Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

[9] Jerry A Hausman, Andrew W Lo, and A Craig MacKinlay. An ordered probit analysis of transaction stock prices. *Journal of financial economics*, 31(3):319–379, 1992.

[10] Gur Huberman and Werner Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275, 2004.

[11] Alan Kraus and Hans R Stoll. Price impacts of block trading on the new york stock exchange. *The Journal of Finance*, 27(3):569–588, 1972.

[12] Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 53(6):1315–1335, 1985.

[13] Charles Lee and Mark J Ready. Inferring trade direction from intraday data. *The Journal of Finance*, 46(2):733–746, 1991.

[14] Andrew E. B. Lim and Poomyos Wimonkittiwat. Dynamic portfolio selection with market impact costs. *Operations Research Letters*, 42(5):299–306, 2014.

[15] Ananth Madhavan and Seymour Smidt. A bayesian model of intraday specialist pricing. *Journal of Financial Economics*, 30(1):99–134, 1991.

[16] James E Matheson and Robert L Winkler. Scoring rules for continuous probability distributions. *Management science*, 22(10):1087–1096, 1976.

[17] Burkart Mönch. Optimal liquidation strategies. *Strategic Trading in Illiquid Markets*, 553:47–78, 2005.

[18] Matthew Spiegel and Avanidhar Subrahmanyam. Asymmetric information and news disclosure rules. *Journal of Financial intermediation*, 9(4):363–403, 2000.

[19] N Torre. Barra market impact model handbook. *BARRA Inc., Berkeley*, 1997.